#2 OIPE

RAW SEQUENCE LISTING   DATE: 07/30/2001
PATENT APPLICATION: US/09/765,272   TIME: 12:02:09

```
Input Set : N:\Crf3\RULE60\09765272.txt
Output Set: N:\CRF3\07302001\I765272.raw
```

```
                   SEQUENCE LISTING
  1 (1) GENERAL INFORMATION:
  5        (i) APPLICANT: Choi et. al.
  9       (ii) TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
 13      (iii) NUMBER OF SEQUENCES: 452
 17       (iv) CORRESPONDENCE ADDRESS:
 19            (A) ADDRESSEE: Human Genome Sciences, Inc.
 21            (B) STREET: 9410 Key West Avenue
 23            (C) CITY: Rockville
 25            (D) STATE: Maryland
 27            (E) COUNTRY: USA
 29            (F) ZIP: 20850
 33        (v) COMPUTER READABLE FORM:
 35            (A) MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 37            (B) COMPUTER: HP Vectra 486/33
 39            (C) OPERATING SYSTEM: MSDOS version 6.2
 41            (D) SOFTWARE: ASCII Text
 45       (vi) CURRENT APPLICATION DATA:
C--> 47       (A) APPLICATION NUMBER: US/09/765,272
C--> 49       (B) FILING DATE: 22-Jan-2001
 51            (C) CLASSIFICATION:
 55      (vii) PRIOR APPLICATION DATA:
 57            (A) APPLICATION NUMBER: 08/961,083
 59            (B) FILING DATE:
 63     (viii) ATTORNEY/AGENT INFORMATION:
 65            (A) NAME: Brookes, A. Anders
 67            (B) REGISTRATION NUMBER: 36,373
 69            (C) REFERENCE/DOCKET NUMBER: PB340P2
C--> 73      (ix) TELECOMMUNICATION INFORMATION:
 75            (A) TELEPHONE: (301) 309-8504
 77            (B) TELEFAX: (301) 309-8512
 85 (2) INFORMATION FOR SEQ ID NO: 1:
 87        (i) SEQUENCE CHARACTERISTICS:
 88            (A) LENGTH: 1999 base pairs
 89            (B) TYPE: nucleic acid
 90            (C) STRANDEDNESS: double
 91            (D) TOPOLOGY: linear
 95       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 97 TAAAATCTAC GACAATAAAA ATCAACTCAT TGCTGACTTG GGTTCTGAAC GCCGCGTCAA          60
 99 TGCCCAAGCT AATGATATTC CCACAGATTT GGTTAAGGCA ATCGTTTCTA TCGAAGACCA         120
101 TCGCTTCTTC GACCACAGGG GGATTGATAC CATCCGTATC CTGGGAGCTT TCTTGCGCAA         180
103 TCTGCAAAGC AATTCCCTCC AAGGTGGATC AACTCTCACC CAACAGTTGA TTAAGTTGAC         240
105 TTACTTTTCA ACTTCGACTT CCGACCAGAC TATTTCTCGT AAGGCTCAGG AAGCTTGGTT         300
107 AGCGATTCAG TTAGAACAAA AAGCAACCAA GCAAGAAATC TTGACCTACT ATATAAATAA         360
109 GGTCTACATG TCTAATGGGA ACTATGGAAT GCAGACAGCA GCTCAAAACT ACTATGGTAA         420
111 AGACCTCAAT AATTTAAGTT TACCTCAGTT AGCCTTGCTG GCTGGAATGC CTCAGGCACC         480
113 AAACCAATAT GACCCCTATT CACATCCAGA AGCAGCCCAA GACCGCCGAA ACTTGGTCTT         540
```

RAW SEQUENCE LISTING   DATE: 07/30/2001
PATENT APPLICATION: US/09/765,272   TIME: 12:02:09

Input Set : N:\Crf3\RULE60\09765272.txt
Output Set: N:\CRF3\07302001\I765272.raw

```
115 ATCTGAAATG AAAAATCAAG GCTACATCTC TGCTGAACAG TATGAGAAAG CAGTCAATAC      600
117 ACCAATTACT GATGGACTAC AAAGTCTCAA ATCAGCAAGT AATTACCCTG CTTACATGGA      660
119 TAATTACCTC AAGGAAGTCA TCAATCAAGT TGAAGAAGAA ACAGGCTATA ACCTACTCAC      720
121 AACTGGGATG GATGTCTACA CAAATGTAGA CCAAGAAGCT CAAAAACATC TGTGGGATAT      780
123 TTACAATACA GACGAATACG TTGCCTATCC AGACGATGAA TTGCAAGTCG CTTCTACCAT      840
125 TGTTGATGTT TCTAACGGTA AAGTCATTGC CCAGCTAGGA GCACGCCATC AGTCAAGTAA      900
127 TGTTTCCTTC GGAATTAACC AAGCAGTAGA AACAAACCGC GACTGGGGAT CAACTATGAA      960
129 ACCGATCACA GACTATGCTC CTGCCTTGGA GTACGGTGTC TACGATTCAA CTGCTACTAT     1020
131 CGTTCACGAT GAGCCCTATA ACTACCCTGG GACAAATACT CCTGTTTATA ACTGGGATAG     1080
133 GGGCTACTTT GGCAACATCA CCTTGCAATA CGCCCTGCAA CAATCGCGAA ACGTCCCAGC     1140
135 CGTGGAAACT CTAAACAAGG TCGGACTCAA CCGCGCCAAG ACTTTCCTAA ATGGTCTAGG     1200
137 AATCGACTAC CCAAGTATTC ACTACTCAAA TGCCATTTCA AGTAACACAA CCGAATCAGA     1260
139 CAAAAAATAT GGAGCAAGTA GTGAAAGAT GGCTGCTGCT TACGCTGCCT TTGCAAATGG     1320
141 TGGAACTTAC TATAAACCAA TGTATATCCA TAAAGTCGTC TTTAGTGATG GGAGTGAAAA     1380
143 AGAGTTCTCT AATGTCGGAA CTCGTGCCAT GAAGGAAACG ACAGCCTATA TGATGACCGA     1440
145 CATGATGAAA ACAGTCTTGA CTTATGGAAC TGGACGAAAT GCCTATCTTG CTTGGCTCCC     1500
147 TCAGGCTGGT AAAACAGGAA CCTCTAACTA TACAGACGAG GAAATTGAAA ACCACATCAA     1560
149 GACCTCTCAA TTTGTAGCAC CTGATGAACT ATTTGCTGGC TATACGCGTA AATATTCAAT     1620
151 GGCTGTATGG ACAGGCTATT CTAACCGTCT GACACCACTT GTAGGCAATG GCCTTACGGT     1680
153 CGCTGCCAAA GTTTACCGCT CTATGATGAC CTACCTGTCT GAAGGAAGCA ATCCAGAAGA     1740
155 TTGGAATATA CCAGAGGGGC TCTACAGAAA TGGAGAATTC GTATTTAAAA ATGGTGCTCG     1800
157 TTCTACGTGG AACTCACCTG CTCCACAACA ACCCCCATCA ACTGAAAGTT CAAGCTCATC     1860
159 ATCAGATAGT TCAACTTCAC AGTCTAGCTC AACCACTCCA AGCACAAATA ATAGTACGAC     1920
161 TACCAATCCT AACAATAATA CGCAACAATC AAATACAACC CCTGATCAAC AAAATCAGAA     1980
163 TCCTCAACCA GCACAACCA                                                  1999
165 (2) INFORMATION FOR SEQ ID NO: 2:
167      (i) SEQUENCE CHARACTERISTICS:
168          (A) LENGTH: 666 amino acids
169          (B) TYPE: amino acid
170          (C) STRANDEDNESS: single
171          (D) TOPOLOGY: linear
173     (ii) MOLECULE TYPE: protein
178     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
180     Lys Ile Tyr Asp Asn Lys Asn Gln Leu Ile Ala Asp Leu Gly Ser Glu
181     1               5                   10                  15
183     Arg Arg Val Asn Ala Gln Ala Asn Asp Ile Pro Thr Asp Leu Val Lys
184                 20                  25                  30
186     Ala Ile Val Ser Ile Glu Asp His Arg Phe Phe Asp His Arg Gly Ile
187             35                  40                  45
189     Asp Thr Ile Arg Ile Leu Gly Ala Phe Leu Arg Asn Leu Gln Ser Asn
190         50                  55                  60
192     Ser Leu Gln Gly Gly Ser Thr Leu Thr Gln Gln Leu Ile Lys Leu Thr
193     65                  70                  75                  80
195     Tyr Phe Ser Thr Ser Thr Ser Asp Gln Thr Ile Ser Arg Lys Ala Gln
196                 85                  90                  95
198     Glu Ala Trp Leu Ala Ile Gln Leu Glu Gln Lys Ala Thr Lys Gln Glu
199                 100                 105                 110
201     Ile Leu Thr Tyr Tyr Ile Asn Lys Val Tyr Met Ser Asn Gly Asn Tyr
202              115                120                  125
```

```
204    Gly Met Gln Thr Ala Ala Gln Asn Tyr Tyr Gly Lys Asp Leu Asn Asn
205        130             135             140
207    Leu Ser Leu Pro Gln Leu Ala Leu Leu Ala Gly Met Pro Gln Ala Pro
208    145             150             155             160
210    Asn Gln Tyr Asp Pro Tyr Ser His Pro Glu Ala Ala Gln Asp Arg Arg
211                165             170             175
213    Asn Leu Val Leu Ser Glu Met Lys Asn Gln Gly Tyr Ile Ser Ala Glu
214        180             185             190
216    Gln Tyr Glu Lys Ala Val Asn Thr Pro Ile Thr Asp Gly Leu Gln Ser
217        195             200             205
219    Leu Lys Ser Ala Ser Asn Tyr Pro Ala Tyr Met Asp Asn Tyr Leu Lys
220        210             215             220
222    Glu Val Ile Asn Gln Val Glu Glu Glu Thr Gly Tyr Asn Leu Leu Thr
223    225             230             235             240
225    Thr Gly Met Asp Val Tyr Thr Asn Val Asp Gln Glu Ala Gln Lys His
226                245             250             255
228    Leu Trp Asp Ile Tyr Asn Thr Asp Glu Tyr Val Ala Tyr Pro Asp Asp
229                260             265             270
231    Glu Leu Gln Val Ala Ser Thr Ile Val Asp Val Ser Asn Gly Lys Val
232        275             280             285
234    Ile Ala Gln Leu Gly Ala Arg His Gln Ser Ser Asn Val Ser Phe Gly
235        290             295             300
237    Ile Asn Gln Ala Val Glu Thr Asn Arg Asp Trp Gly Ser Thr Met Lys
238    305             310             315             320
240    Pro Ile Thr Asp Tyr Ala Pro Ala Leu Glu Tyr Gly Val Tyr Asp Ser
241                325             330             335
243    Thr Ala Thr Ile Val His Asp Glu Pro Tyr Asn Tyr Pro Gly Thr Asn
244                340             345             350
246    Thr Pro Val Tyr Asn Trp Asp Arg Gly Tyr Phe Gly Asn Ile Thr Leu
247        355             360             365
249    Gln Tyr Ala Leu Gln Gln Ser Arg Asn Val Pro Ala Val Glu Thr Leu
250        370             375             380
252    Asn Lys Val Gly Leu Asn Arg Ala Lys Thr Phe Leu Asn Gly Leu Gly
253    385             390             395             400
255    Ile Asp Tyr Pro Ser Ile His Tyr Ser Asn Ala Ile Ser Ser Asn Thr
256                405             410             415
258    Thr Glu Ser Asp Lys Lys Tyr Gly Ala Ser Ser Glu Lys Met Ala Ala
259                420             425             430
261    Ala Tyr Ala Ala Phe Ala Asn Gly Gly Thr Tyr Tyr Lys Pro Met Tyr
262        435             440             445
264    Ile His Lys Val Val Phe Ser Asp Gly Ser Glu Lys Glu Phe Ser Asn
265        450             455             460
267    Val Gly Thr Arg Ala Met Lys Glu Thr Thr Ala Tyr Met Met Thr Asp
268    465             470             475             480
270    Met Met Lys Thr Val Leu Thr Tyr Gly Thr Gly Arg Asn Ala Tyr Leu
271                485             490             495
273    Ala Trp Leu Pro Gln Ala Gly Lys Thr Gly Thr Ser Asn Tyr Thr Asp
274                500             505             510
276    Glu Glu Ile Glu Asn His Ile Lys Thr Ser Gln Phe Val Ala Pro Asp
```

RAW SEQUENCE LISTING   DATE: 07/30/2001
PATENT APPLICATION: **US/09/765,272**   TIME: 12:02:09

Input Set : **N:\Crf3\RULE60\09765272.txt**
Output Set: **N:\CRF3\07302001\I765272.raw**

```
277           515              520                  525
279     Glu Leu Phe Ala Gly Tyr Thr Arg Lys Tyr Ser Met Ala Val Trp Thr
280         530              535              540
282     Gly Tyr Ser Asn Arg Leu Thr Pro Leu Val Gly Asn Gly Leu Thr Val
283     545              550              555              560
285     Ala Ala Lys Val Tyr Arg Ser Met Met Thr Tyr Leu Ser Glu Gly Ser
286                 565                  570                  575
288     Asn Pro Glu Asp Trp Asn Ile Pro Glu Gly Leu Tyr Arg Asn Gly Glu
289             580              585              590
291     Phe Val Phe Lys Asn Gly Ala Arg Ser Thr Trp Asn Ser Pro Ala Pro
292         595              600              605
294     Gln Gln Pro Pro Ser Thr Glu Ser Ser Ser Ser Ser Ser Asp Ser Ser
295         610              615              620
297     Thr Ser Gln Ser Ser Ser Thr Thr Pro Ser Thr Asn Asn Ser Thr Thr
298     625              630              635              640
300     Thr Asn Pro Asn Asn Asn Thr Gln Gln Ser Asn Thr Thr Pro Asp Gln
301             645              650              655
303     Gln Asn Gln Asn Pro Gln Pro Ala Gln Pro
304             660              665
```

306 (2) INFORMATION FOR SEQ ID NO: 3:
308   (i) SEQUENCE CHARACTERISTICS:
309   (A) LENGTH: 1714 base pairs
310   (B) TYPE: nucleic acid
311   (C) STRANDEDNESS: double
312   (D) TOPOLOGY: linear
316   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```
318 AAATTACAAT ACGGACTATG AATTGACCTC TGGAGAAAAA TTACCTCTTC CTAAAGAGAT     60
320 TTCAGGTTAC ACTTATATTG GATATATCAA AGAGGGAAAA ACGACTTCTG AGTCTGAAGT    120
322 AAGTAATCAA AAGAGTTCAG TTGCCACTCC TACAAAACAA CAAAAGGTGG ATTATAATGT    180
324 TACACCGAAT TTTGTAGACC ATCCATCAAC AGTACAAGCT ATTCAGGAAC AAACACCTGT    240
326 TTCTTCAACT AAGCCGACAG AAGTTCAAGT AGTTGAAAAA CCTTTCTCTA CTGAATTAAT    300
328 CAATCCAAGA AAAGAAGAGA AACAATCTTC AGATTCTCAA GAACAATTAG CCGAACATAA    360
330 GAATCTAGAA ACGAAGAAAG AGGAGAAGAT TTCTCCAAAA GAAAAGACTG GGGTAAATAC    420
332 ATTAAATCCA CAGGATGAAG TTTTATCAGG TCAATTGAAC AAACCTGAAC TCTTATATCG    480
334 TGAGGAAACT ATGGAGACAA AAATAGATTT TCAAGAAGAA ATTCAAGAAA ATCCTGATTT    540
336 AGCTGAAGGA ACTGTAAGAG TAAAACAAGA AGGTAAATTA GGTAAGAAAG TTGAAATCGT    600
338 CAGAATATTC TCTGTAAACA AGGAAGAAGT TTCGCGAGAA ATTGTTTCAA CTTCAACGAC    660
340 TGCGCCTAGT CCAAGAATAG TCGAAAAAGG TACTAAAAAA ACTCAAGTTA TAAAGGAACA    720
342 ACCTGAGACT GGTGTAGAAC ATAAGGACGT ACAGTCTGGA GCTATTGTTG AACCCGCAAT    780
344 TCAGCCTGAG TTGCCCGAAG CTGTAGTAAG TGACAAAGGC GAACCAGAAG TTCAACCTAC    840
346 ATTACCCGAA GCAGTTGTGA CCGACAAAGG TGAGACTGAG GTTCAACCAG AGTCGCCAGA    900
348 TACTGTGGTA AGTGATAAAG GTGAACCAGA GCAGGTAGCA CCGCTTCCAG AATATAAGGG    960
350 TAATATTGAG CAAGTAAAAC CTGAAACTCC GGTTGAGAAG ACCAAAGAAC AAGGTCCAGA   1020
352 AAAAACTGAA GAAGTTCCAG TAAAACCAAC AGAAGAAACA CCAGTAAATC CAAATGAAGG   1080
354 TACTACAGAA GGAACCTCAA TTCAAGAAGC AGAAAATCCA GTTCAACCTG CAGAAGAATC   1140
356 AACAACGAAT TCAGAGAAAG TATCACCAGA TACATCTAGC AAAAATACTG GGAAGTGTC    1200
358 CAGTAATCCT AGTGATTCGA CAACCTCAGT TGGAGAATCA AATAAACCAG AACATAATGA   1260
360 CTCTAAAAAT GAAAATTCAG AAAAAACTGT AGAAGAAGTT CCAGTAAATC CAAATGAAGG   1320
362 CACAGTAGAA GGTACCTCAA ATCAAGAAAC AGAAAAACCA GTTCAACCTG CAGAAGAAAC   1380
```

RAW SEQUENCE LISTING   DATE: 07/30/2001
PATENT APPLICATION: US/09/765,272   TIME: 12:02:09

Input Set : **N:\Crf3\RULE60\09765272.txt**
Output Set: **N:\CRF3\07302001\I765272.raw**

```
364 ACAAACAAAC TCTGGGAAAA TAGCTAACGA AAATACTGGA GAAGTATCCA ATAAACCTAG     1440
366 TGATTCAAAA CCACCAGTTG AAGAATCAAA TCAACCAGAA AAAAACGGAA CTGCAACAAA     1500
368 ACCAGAAAAT TCAGGTAATA CAACATCAGA GAATGGACAA ACAGAACCAG AACCATCAAA     1560
370 CGGAAATTCA ACTGAGGATG TTTCAACCGA ATCAAACACA TCCAATTCAA ATGGAAACGA     1620
372 AGAAATTAAA CAAGAAAATG AACTAGACCC TGATAAAAAG GTAGAAGAAC CAGAGAAAAC     1680
374 ACTTGAATTA AGAAATGTTT CCGACCTAGA GTTA                                1714
376 (2) INFORMATION FOR SEQ ID NO: 4:
378     (i) SEQUENCE CHARACTERISTICS:
379         (A) LENGTH: 571 amino acids
380         (B) TYPE: amino acid
381         (C) STRANDEDNESS: single
382         (D) TOPOLOGY: linear
384    (ii) MOLECULE TYPE: protein
389    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
391    Asn Tyr Asn Thr Asp Tyr Glu Leu Thr Ser Gly Glu Lys Leu Pro Leu
392    1               5                   10                  15
394    Pro Lys Glu Ile Ser Gly Tyr Thr Tyr Ile Gly Tyr Ile Lys Glu Gly
395                   20                  25                  30
397    Lys Thr Thr Ser Glu Ser Glu Val Ser Asn Gln Lys Ser Ser Val Ala
398               35                  40                  45
400    Thr Pro Thr Lys Gln Gln Lys Val Asp Tyr Asn Val Thr Pro Asn Phe
401           50                  55                  60
403    Val Asp His Pro Ser Thr Val Gln Ala Ile Gln Glu Gln Thr Pro Val
404    65                  70                  75                  80
406    Ser Ser Thr Lys Pro Thr Glu Val Gln Val Val Glu Lys Pro Phe Ser
407                   85                  90                  95
409    Thr Glu Leu Ile Asn Pro Arg Lys Glu Glu Lys Gln Ser Ser Asp Ser
410               100                 105                 110
412    Gln Glu Gln Leu Ala Glu His Lys Asn Leu Glu Thr Lys Lys Glu Glu
413               115                 120                 125
415    Lys Ile Ser Pro Lys Glu Lys Thr Gly Val Asn Thr Leu Asn Pro Gln
416           130                 135                 140
418    Asp Glu Val Leu Ser Gly Gln Leu Asn Lys Pro Glu Leu Leu Tyr Arg
419    145                 150                 155                 160
421    Glu Glu Thr Met Glu Thr Lys Ile Asp Phe Gln Glu Glu Ile Gln Glu
422                   165                 170                 175
424    Asn Pro Asp Leu Ala Glu Gly Thr Val Arg Val Lys Gln Glu Gly Lys
425               180                 185                 190
427    Leu Gly Lys Lys Val Glu Ile Val Arg Ile Phe Ser Val Asn Lys Glu
428           195                 200                 205
430    Glu Val Ser Arg Glu Ile Val Ser Thr Ser Thr Thr Ala Pro Ser Pro
431           210                 215                 220
433    Arg Ile Val Glu Lys Gly Thr Lys Lys Thr Gln Val Ile Lys Glu Gln
434    225                 230                 235                 240
436    Pro Glu Thr Gly Val Glu His Lys Asp Val Gln Ser Gly Ala Ile Val
437                   245                 250                 255
439    Glu Pro Ala Ile Gln Pro Glu Leu Pro Glu Ala Val Val Ser Asp Lys
440               260                 265                 270
442    Gly Glu Pro Glu Val Gln Pro Thr Leu Pro Glu Ala Val Val Thr Asp
```

VERIFICATION SUMMARY
PATENT APPLICATION: **US/09/765,272**

DATE: 07/30/2001
TIME: 12:02:10

Input Set : **N:\Crf3\RULE60\09765272.txt**
Output Set: **N:\CRF3\07302001\I765272.raw**

L:47 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:49 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:73 M:220 C: Keyword misspelled or invalid format, [(ix) TELECOMMUNICATION INFORMATION:]
L:853 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10
L:2649 M:111 C: (47) String data converted to upper case,
M:111 Repeated in SeqNo=41
L:2885 M:111 C: (47) String data converted to upper case,
M:111 Repeated in SeqNo=45
L:2927 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:46
L:2984 M:111 C: (47) String data converted to upper case,
M:111 Repeated in SeqNo=47
L:3320 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:52
L:3641 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:56
L:3825 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:58
L:3828 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:58
L:4317 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:66
L:4395 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:66
L:4966 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:74
L:5102 M:111 C: (47) String data converted to upper case,
M:111 Repeated in SeqNo=75
L:5172 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:76
L:5365 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:80
L:5398 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:80
L:5533 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:82
L:7004 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:106
L:7007 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:106
L:7019 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:106
L:7034 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:106
L:7075 M:111 C: (47) String data converted to upper case,
M:111 Repeated in SeqNo=107
L:7616 M:111 C: (47) String data converted to upper case,
L:7764 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:118
L:7944 M:111 C: (47) String data converted to upper case,
L:8040 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:120
L:10220 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:160
L:10343 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:162
L:10809 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:172
L:10812 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:172
L:10815 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:172
L:11039 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:176
L:11932 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:194
L:14181 M:111 C: (47) String data converted to upper case,
L:17498 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:452